

잔류가스 분석기(RGA)와 인공지능 모델링을 이용한 모니터링 시스템 개발

이지수* · 김승훈** · 김경수*** · 송효종***† · 박상훈**** · 고득훈**** · 이봉재****

*명지대학교 산업경영공학과, **명지대학교 전자공학과,
†명지대학교 환경에너지공학과, *(재)한국화학융합시험연구원

Development of Monitoring System Using Residual Gas Analyzer (RGA) and Artificial Intelligence Modeling

Ji Soo Lee*, Song Hun Kim**, Gyeong Su Kim***, Hyo Jong Song***†,
Sang-Hoon Park****, Deuk-Hoon Goh**** and Bong-Jae Lee****

*Department of Industrial Management of Engineering, Myongji University,

**Department of Electronic Engineering, Myongji University,

***†Department of Environmental Energy and Engineering, Myongji University,

****Korea Testing & Research Institute

ABSTRACT

This study aims to talk about the necessity of solving the PFC gas emission problem raised by the recent development of the semiconductor industry and the remote plasma source method monitoring system used in the semiconductor industry. The 'monitoring system' means that the researchers applied machine learning to the existing monitoring technology and modeled it. In the process of this study, Residual Gas Analyzer monitoring technology and linear regression model were used. Through this model, the researchers identified emissions of at least 12700mg CO₂ to 75800mg CO₂ with values ranging from ion current 0.6A to 1.7A, and expect that the 'monitoring system' will contribute to the effective calculation of greenhouse gas emissions in the semiconductor industry in the future.

Key Words : Remote Plasma Source(RPS), Residual Gas Analysis(RGA), Linear Regression(LR)

1. 서 론

최근 한국 반도체 산업의 글로벌 수출시장 점유율이 감소하는 추세지만, 한국의 대표적 수출 품목인 메모리 반도체가 2018년 29.1%로, 세계 수출시장 점유율 1위를 수년간 유지해왔다는 점[1] 과, 정부가 최근 '반도체 메가 클러스터' 계획을 통해 국내 반도체 산업 제조공정의 경쟁력을 높이고 있다는 점에서 반도체 디스플레이 산업이 한

국의 성장 가능성이 가장 큰 주력 산업임을 알 수 있다.

특히 환경보전의 관점에서 파리협정 1.5도 목표 실현을 위한 한국의 2030 NDC와 2050 탄소중립 목표[2]를 위해서, 반도체 디스플레이 주력 산업의 온실가스 감축은 꼭 해결해야 할 과제이다.

그리고 온실가스 중 반도체 산업의 PFC, HFCs, SF₆, NF₃ 가스는 Table 1과 같이 지구온난화 지수(Global Warming Potential, GWP)가 CO₂에 비해 수천 배 가량 높고, 대기 잔류 시간 또한 길기에[3], 최근 반도체 공정 내 PFCs 가스 배출량 저감을 위한 연구가 활발히 이루어지고 있는 추세이다.

†E-mail: hjsong@mju.ac.kr

Table 1. 주요 PFC 가스의 대기 잔류 시간 및 GWP100 [4]

Greenhouse gases	Atmospheric lifetime(year)	GWP100
CO ₂	-	1
CH ₄	12	25
N ₂ O	109	298
CF ₄	50000	6500
C ₃ F ₈	2600	8830
C ₄ F ₁₀	2600	8860
C ₂ F ₆	10000	12200
CHF ₃	270	14800
NF ₃	740	17200
SF ₆	3200	22800

그 예로 최신 반도체 및 디스플레이 공정의 경우 주로 높은 분해 효율과 정밀한 패턴 형성 덕분에, CF₄ 원격 플라즈마 (Remote plasma source, RPS)가 사용되고 있다. 이때, CF₄ 원격 플라즈마는 반도체 제조 공정 간 식각(Etch) 과정에서 CF₄ 가스를 활성화시켜 사용되며[5], 층 간의 구조를 정의하거나 원하는 패턴을 형성하는 데 활용된다.

본 연구는 최근 반도체 산업의 발전과 함께 제기되는 CF₄ 원격 플라즈마를 이용한 식각(Etch) 과정 내 PFCs가스에 대한 모니터링 방법과 배출량 산정의 필요성에 대해 이야기하고자 한다. 또한, 과거 논문들에서 인자들 간의 선형성을 중점적으로 파악했고 기존에 정량화된 연구가 미비했던 점을 고려해 정확한 배출량 산정을 위한 식을 수립하고 정량적인 배출량을 파악하고자 한다.

본 논문에서 언급될 ‘모니터링 시스템’이란, 기존 RGA 실시간 모니터링 기술에 머신 러닝을 접목하여, 모델링화 시킨 것을 의미한다.

이때, 잔류가스 분석기인 Residual Gas Analyzer (RGA)[6]와 해당 분석기를 정량평가 하는 FTIR을 통한 실시간 모니터링 정량분석기술과 선형회귀(Linear Regression) 모델링이 연구 과정에서 사용되었다.

RGA는 진공시스템 내 잔류 가스를 측정하거나, 공정 시스템 내 반응가스 및 생성가스의 변화를 관찰하는 기술이며, FTIR은 잔류가스의 농도를 측정하여 RGA의 분석 결과에 대한 정량평가를 수행하는데 활용되었다.

본 실험에서는 해당 기술을 이용해 배가스의 CF₄ 농도(ppm)를 측정하기 위해 FTIR을 활용했다. 이후 6차 IPCC Assessment report을 참고하여 Y=0.06XZ라는 배출량 산정식을 통해 CF₄ 배출량을 정량화 했다.

따라서 정량화된 RGA 데이터를 기반으로 공정 데이터

의 선형 회귀를 통해 도출된 결과를 토대로 실험 조건에서 이온 전류를 0.6A에서 1.7A로 설정한 결과, 2월과 4월에 진행된 실험조건에서 정량적인 CF₄ 배출량이 12,700mg CO₂에서 75,800mg CO₂까지 측정되었다.

본 연구진은 정량적인 CF₄ 배출량 측정 과정을 통한 ‘실시간 모니터링 시스템’의 구축을 목표로 하며, 더 나아가 해당 시스템이 향후 반도체 공정 상의 효율적인 온실가스 배출량 산정에 일조하기를 바란다.

2. 모델링 과정

본 연구의 정량적인 배출량 산정을 위하여 다음과 같은 순서로, 각 변수들(ion current와 ppm)간의 선형성과 선형회귀 모델을 통한 배출량 산정식의 도출하였다.

2.1 FTIR 농도(ppm) 전처리 과정

먼저 연구 과정에서 “한국화학융합시험연구원”의 24년 2월, 4월 실험 데이터를 사용하였으며, CF₄ Ion Current에 따른 CF₄ 농도 변화를 추적하고자 FTIR로 측정된 CF₄ 시계열 데이터를 분석하였다. 이 때, 시계열 데이터란 시간 순서에 따라 관측된 데이터로 CF₄ 농도(ppm)이 이것에 해당된다. 따라서 실험조건은 아래의 Table 2와 같다.

[‘챔버 내부 압력=25 mTorr’, ‘플라즈마 세기=1 kW’, ‘추가 가스(Ar) 유입량=500 sccm’, ‘측정 시간=3분’]

Table 2. Experimental conditions.

EXP_24_02,04 실험조건							
공정 및 가스	고정변수			측정 유량	미지 유량	측정 시간	
	챔버 내부 압력	플라즈마 세기	추가 가스				
Etching (식각) CF ₄	25 mTorr	1KW	Ar 500	90	100,	3분	
				120			
				150			180,
				180			200
				210			1분 휴식
240							

Table 3. Quantitative value between FTIR(ppm) and flow rate

Flow Rate	FTIR(ppm)
90	2214.50
120	3132.53
150	4046.87
180	4861.84
210	5653.91
240	6573.20

유량에 따른 CF₄ 농도의 정량적 변화를 추적하고자, 위의 실험 조건의 2월, 4월 실험 데이터를 Table 3과 같이 정리하였다.

이 표를 통해 유량과 FTIR 측정 농도(ppm) 간의 비례 관계를 확인할 수 있었고, 유량 조절이 CF₄ 농도에 선형적으로 영향을 미치는지를 정량적으로 파악 할 수 있었다. 이에 더해, 4월의 실험에서는 플라즈마를 켜고 끄는 두 가지 상황에서 실험을 진행하였다.

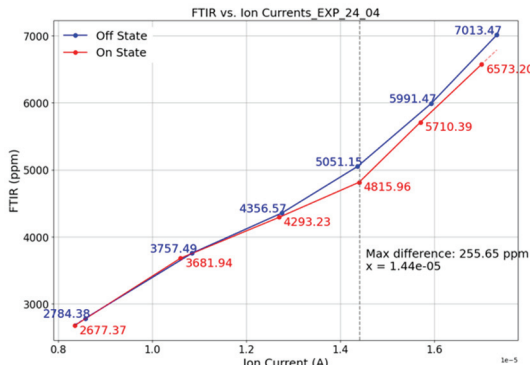


Fig. 1. Comparison of emission concentrations during plasma on off.

Fig 1은 4월에 시행된 실험 결과를 보여주며, 이를 통해 플라즈마가 CF₄ 제거에 미치는 영향을 확인하고자 했다. 실험 데이터를 분석하여 플라즈마를 켜고 끄는 때의 CF₄ 농도 차이를 관찰했다. 이 농도 차이는 이온전류가 1.44A일때 최대 255.65 ppm로 나타났다.

이러한 결과는 RPS 플라즈마가 CF₄를 효과적으로 제거하는데 도움이 되는 것으로 나타났고 따라서 이 연구는 RPS 플라즈마가 온실가스인 CF₄를 제거하는데 유망한 가능성을 시사한다.

2.2 FTIR 농도(ppm)와 이온전류의 선형성

본 연구는 CF₄ 이온전류와 CF₄ 농도(ppm)의 선형성을 파악함으로써, 향후 실험 또는 시스템 설계 시 CF₄ 제거 효과를 예측하고 조절하고자 했다. 따라서 2월, 4월 실험 데이터를 Fig.2 그래프로 표현했다.

Fig. 2의 결과에 따르면, 이온전류와 CF₄ 농도 사이에는 선형적인 관계가 파악됨을 알 수 있었다. 따라서 이온전류의 증가에 따라 CF₄ 농도도 증가하는 경향을 보이는 결론을 얻었다. 이러한 결과는 두 변수 간의 직접적인 상관성을 보여주는 것임을 확인할 수 있었다.

이를 통해 이온전류의 변화에 따른 CF₄ 농도의 예측값을 얻어, 향후 3.2의 실험 계획 및 시스템 설계에 활용할 것이다.

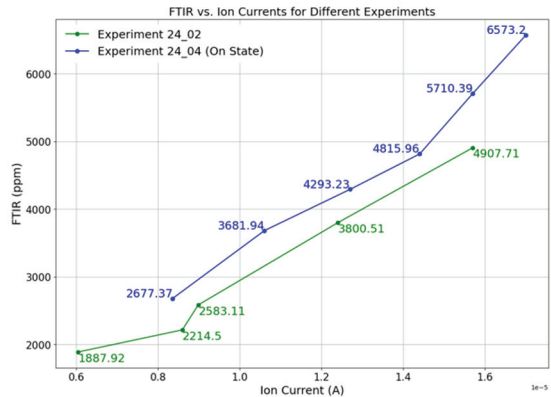


Fig. 2. Peak compared to ion current in February and April.

2.3 반도체 유해가스 배출량 산정

반도체 유해가스 CO₂ 환산 배출량 도출식 = **KXZ**

$$F(X) = (1 - \text{CF}_4 \text{ 가스 잔류비율}) \times (1 - \text{CF}_4 \text{ 가스사용률}) \times X \times Z \times \frac{\text{오염물질(g)}}{22.4(\text{L})} \times \text{반응시간(min)} \times \text{오염물질 GWP} \quad (1)$$

$$K = (1 - \text{CF}_4 \text{ 가스 잔류비율}) \times (1 - \text{CF}_4 \text{ 가스사용률}) \times 8\text{g}/22.4(\text{L}) \times 3\text{min} \times 6500 \text{ CO}_2 = 0.06 [4]$$

K는 반도체 유해가스 CO₂ 환산 배출량 도출식에서 사용되는 상수이다. 이는 주어진 환산식에서 사용되는 여러 인자들의 값을 결합하여 최종적으로 CO₂ 배출량을 계산하는 데 필요한 값이다.

X는 유입 유량을 나타내는 변수로, 이는 시간당 장치로 유입되는 가스의 양을 나타낸다. 단위는 mL/min으로 표시되며, 이는 시간당 mL로 표시된 가스의 양이다.

Z는 유해가스의 농도를 나타내는 변수로, ppm(백만분율)으로 표시되며, 대기 중의 가스 중 해당 유해가스의 부피가 차지하는 비율을 나타낸다.

3. 결과 분석

3.1 데이터 분석

위 2.3의 반도체 CF₄ 유해가스 배출량산정식을 통하여 참조하여 24년 2월과 4월에 시행된 실험에 대해 CO₂ 대비 정량적 배출량을 Table 4 표로 표현하였다.

Table 4. Quantitative emission compared to CO₂ (24/02, 24/04)

Ion Currents (×10 ⁻⁵)	Emission estimate	
	February	April
0.8	12900	14800
1.0	23900	23900
1.2	33200	35500
1.4	44500	50200
1.6	-	79300

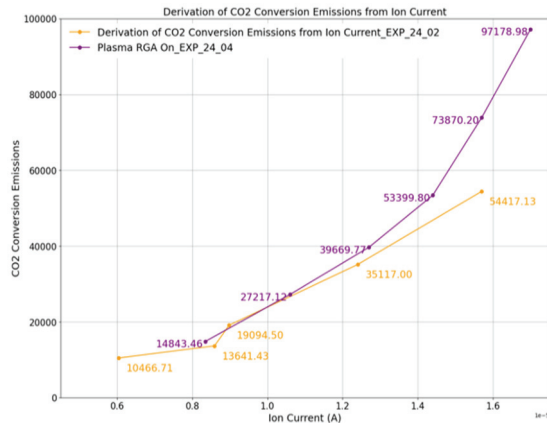


Fig. 3. Quantitative calculation of semiconductor emissions through the calculation formula for February and April.

Fig 3은 실제 데이터에 대한 2.3의 산정식 결과를 보여 주며, 향후 기업에서 같은 실험 조건의 RGA 실시간 모니터링에서 얻은 ion current와 비교하여 CO₂ 대비 배출량을 파악할 수 있도록 그래프로 정리했다.

이 때, ion current는 유량과 농도(ppm)와의 상관성을 띄기 때문에 2.3의 산정식 중 X와 Z에 기여하는 값이며 본 결과를 통해, 유해가스의 CO₂ 환산 배출량과 이온 전류 간의 상관성을 파악할 수 있었다.

또한 위와 같은 고찰은 CO₂ 환산 배출량 예측 모델의 출발점으로 활용하였다. 본 연구에서는 선형회귀 모델링과 랜덤 포레스트 모델링 중 CO₂ 환산 배출량 예측에 적합한 모델을 찾고자 한다.

3.2 Linear regression 성능검정

MSE (Mean Squared Error)와 R²(결정계수, Coefficient of Determination)은 회귀 분석의 평가지표 중 하나로, 각각 예측실제 값 간의 차이를 제공하여 평균한 값과 모델이 종속 변수의 변동성에 대한 값을 의미한다. 이 때 R²값이 1에 가까울수록 모델 적합도가 높다고 할 수 있다.

선형회귀[7]의 성능검정을 위해 비교군으로 높은 예측

성능과 과적합(Overfitting)을 방지하는 특성을 가진 랜덤 포레스트(Random Forest) 방식[8]을 실험값에 적용하고 선형회귀 방식의 모델링과 비교하는 과정을 수행했다. 이때 모델의 성능 평가 결과 MSE_{RF} < MSE_{LR} 라는 점에서, 모델의 성능 평가 결과 랜덤 포레스트 모델의 성능이 우수함을 알 수 있었다[9]. 본 결과를 바탕으로 선형회귀 보다 랜덤 포레스트 방식이 성능적인 우위에 있음을 파악했다.

Table 5. Performance Evaluation Metrics for Linear Regression.

	Experimental Models	Measurements	
		MSE	R ²
Results	1) Random forest	6423446	0.985
	2) Linear Regression	23836524	0.946

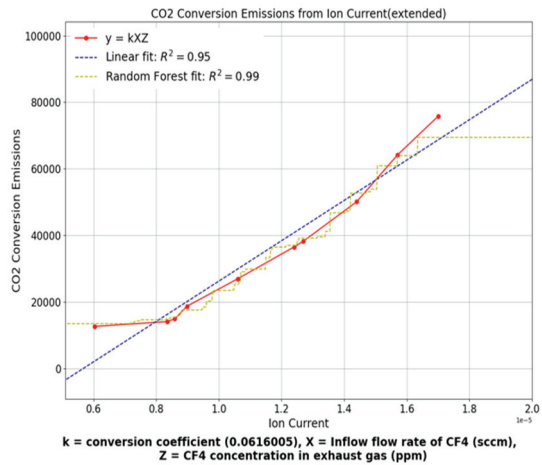


Fig. 4. Comparison of model fit between random forest and linear regression.

그러나 Fig 4에서 확인할 수 있듯이 랜덤 포레스트 방식의 회귀 모델은 학습한 데이터 너머의 예측 값들이 거의 동일하게 예측하는 양상을 보였으며 이는 과적합에 해당하는 것으로 판단되었다. 따라서 기존 학습 범위의 데이터에 대한 예측은 랜덤 포레스트 모델링의 성능이 비교적 우수하다고 할 수 있겠으나, 머신러닝 모델의 일반화 성능을 고려하였을 때에는 선형회귀 모델이 더 타당하다고 판단되었다.

또한 이러한 결과를 고찰하였을 때 머신러닝 모델의 학습 데이터셋의 양과 다양성이 충분히 확보되지 않았음을 알 수 있으며, 이는 랜덤 포레스트 모델과 같은 의사결정 트리 모델이 알고리즘의 특성 상 실험 값 예측 시에 과적

합의 위험에 더 취약하다는 점을 확인할 수 있었다[10]. 따라서 향후 연구를 위한 제언점으로 모니터링 시스템의 구축 및 고도화에 있어 의사결정 트리에 기반한 알고리즘이 높은 성능을 나타낼 수 있음에도 과적합 문제가 빈번하게 발생할 수 있음에 유의하여야 함을 제시한다.

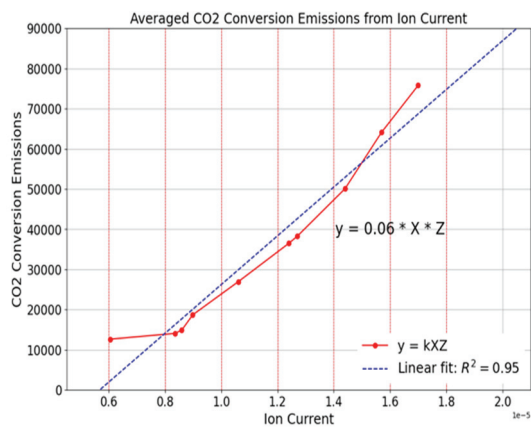
4. 결 론

본 연구는 모니터링 시스템을 위한 머신러닝 알고리즘의 선택에 있어 선형회귀와 같이 단순하지만 연속적인 값에 대하여 예측 문제에 효과적으로 대응할 수 있는 알고리즘[11]을 고려했다.

결과적으로 연구에는 일반화 성능 반복적으로 측정된 연속형 값인 점을 고려할 때, 선형 회귀 모델이 데이터 분석에 더 적합하다는 결론을 도출할 수 있었다. 이는 연속적인 값에 대한 예측 문제를 해결하는 데 있어서 선형 회귀 모델이 더 적절하다는 것을 시사한다.

Table 6. Quantitative emission compared to CO₂

Ion Currents (×10 ⁻⁵)	Emission estimate	
	평균치 기반 예측값	선형회귀 모델링
0.8	13900	14100
1.0	23900	26300
1.2	34400	38400
1.4	47300	50500
1.6	66800	62600
1.8	-	74800



k = conversion coefficient (0.0616005), X = Inflow flow rate of CF₄ (sccm), Z = CF₄ concentration in exhaust gas (ppm)

Fig. 5. Expression of emission relative to ion current in linear regression(y=0.06XZ).

실험 데이터를 분석한 결과, 이온 전류가 0.6인 경우 CF₄의 배출량은 12,700mg CO₂임을 확인할 수 있었으며, 이 값이 1.7로 증가할 경우 CF₄의 배출량은 75,800mg CO₂으로 증가함을 알 수 있었다. 또한, 이를 기반으로 예측된 2.0의 이온 값은 86900mg CO₂로 더 많은 CF₄ 배출량을 나타내는 것을 알 수 있었다.

탄소중립 달성 및 ESG경영 달성을 위해, 각 기업의 PFCs 가스 배출량 산정은 필수적인 요소이다. 하지만 모든 배가스 배출량 중, 특정 PFCs 가스에 대한 실시간 배출량 산정은 여러 측정 장비가 필요하다는 문제점이 존재한다.

이 모니터링 시스템에서는 'LR(Linear Regression) 머신러닝' 과정에서 '반도체 유해가스 CO₂ 환산 배출량 도출식 = KXZ'을 인자로 사용한다. 이는 기업이 이산화탄소로 환산된 PFCs 가스의 배출량을 예측하는 데 기여할 수 있다.

그러나 이 공식은 GWP를 파악하지 못하는 다른 PFCs에 대해서는 적용이 제한된다는 한계가 있다. 그러나 만약 미지의 가스에 대한 GWP 환산식을 구할 수 있다면, 본 연구에서 고안한 '반도체 유해가스 CO₂ 환산 배출량 도출식 = 0.06XZ'는 모든 PFCs 가스에 대하여 적용될 수 있다.

이러한 연구 결과는 향후 반도체 산업에서 발생하는 유해 가스의 배출량을 정확히 도출하고 관리하는 데 큰 도움이 될 것으로 기대된다.

감사의 글

이 연구는 2023년도 산업통상자원부 및 한국산업기술기획평가원(KEIT) 연구비 지원에 의한 연구(RS-2023-00267529) 및 2024년도 부처협업형 반도체전공트랙 사업을 통해 한국산업기술진흥원(G02P1880005502)의 지원을 받아 수행된 연구임.

참고문헌

- H.G.Jeong, "Import and Export Structure and Global Phase Analysis of the Korean Semiconductor Industry." *World Economy Today* 24.2 (2024): 1-30.
- Ministry of Foreign Affairs, "climate change negotiations", https://www.mofa.go.kr/www/wpge/m_20150/contents.do, (2024.04.01.)
- S.H. Han, H.W. Park, T.H. Kim, and D.W. Park, "Large scale treatment of perfluorochemicals using a thermal plasma scrubber." *Clean Technology* 17, no. 3 (2011): 250-258.
- Intergovernmental Panel on Climate Change (IPCC).

- "IPCC 6th Evaluation Report Comprehensive Report." Panel's 58th Session held in Interlaken, Switzerland (2023).
5. J.Y. Ryu, "A Study on Decomposition of Perfluorocompounds (NF₃, CF₄, SF₆) Emitted from Semiconductor Process using Plasma." *Journal of Korea Society of Waste Management* 35.3 (2018): 250-257.
 6. C.J. Park, "Operating Principle of Residual Gas Analyzer." *Applied Science and Convergence Technology* 17.4 (2008): 262-269.
 7. S.Y. Lee and Y.J. Huh, "A Study on Prediction Model of Scaffold Pore Size Using Machine Learning." *Journal of the Semiconductor & Display Technology* 18.4 (2019): 49.
 8. K.H. Kang and H.J. Park, "Study on the Effect of Training Data Sampling Strategy on the Accuracy of the Landslide Susceptibility Analysis Using Random Forest Method." *Economic and Environmental Geology* 52.2 (2019): 199-212.
 9. G.W. Cha and W.H. Hong, "Development of an Predictive Model via Decision Tree-Based Algorithms for Forecasting Demolition Waste Generation." *Journal of the Architectural Institute of Korea (JAIK)* 39.3 (2023): 179-187.
 10. S.C. Ha. "A Research on the Selection Algorithm in Trading Strategies using Machine Learning Techniques." Kookmin University, Seoul (2016): 1-7.
 11. D.I. Kim, C.S. Park, J.G. Baek, and S.S. Kim, "Multi-objective Genetic Algorithm for Variable Selection in Linear Regression Model and Application." *Journal of the Korea society for simulation* 18.4 (2009): 137-148.
-
- 접수일: 2024년 6월 7일, 심사일: 2024년 6월 17일,
 게재확정일: 2024년 6월 21일